Introduction to the LANL HPC Environment

LANL CESM Tutorial
26 April 2011

Sara Rauscher, Keeley Costigan

Adapted from LANL HPC Online Guides

Connecting to Mapache

To reach Turquoise HPC clusters (from inside LANL or outside), you need three things:

- an approved and established HPC account
- your LANL Cryptocard
- ssh software on your workstation to get through the ssh proxy firewall system, wtrw.lanl.gov.

ssh -t -X wtrw.lanl.gov ssh mp-fe1.lanl.gov *or*

ssh -t -X wtrw.lanl.gov ssh mp-fe2.lanl.gov

Use your cryptocard to generate a password

More information here (cryptocard needed to access these pages):

http://hpc.lanl.gov/?q=getting_started
http://hpc.lanl.gov/turquoise_access
http://hpc.lanl.gov/mapache_home

Logging into Mapache from Encantado Classroom Terminals

- Log on to local PC
  - Username is PC name; encX (where X is 01 to 15)
  - Password is on white board
- Bring up XWinServer
- Bring up Putty
  - Enter wtrw.lanl.gov as the machine to which you will ssh
  - On the left panel of the Putty window, click on the + next to SSH
  - Click on X11 below SSH
  - Check the box in the right panel that allows X11 forwarding
  - Open
  - Login to wtrw as your username/moniker
  - Use your crypto card for the password
- SSH to Mapache
  - ssh mp-fe1.lanl.gov or ssh mp-fe2.lanl.gov
- wtrw re-authentication does not appear to work in Putty and you may need to restart it after four hours

User Environment

- The user interfaces to the operating system are called shells. The shells are programs that interpret the commands you enter, run the programs you have asked for, and send the results to your screen.

- Default login shell on all LANL ASCI systems is **tcsh**. *Trusted csh*,a superset of csh, with everything csh has and more.
  - tcsh is the only interactive shell supported.
  - Use other shells at your own risk.
  - Inconsistent results may result if any other shell is used.

- One of the really nice features is command-line editing using the arrow keys:
  - Use the "up" arrow to capture previous commands.
  - Use the "left" arrow to back over the characters in a previous command.
  - Any typing you do will *insert*.

User Space on Mapache: we'll be using the areas in red

- /users/username/
  - 500 MB (small space)
  - Recommended use: executables and small source code. But practically you should use your project space
  - Backed up regularly
- /usr/projects/projectname/username
  - Limited allocation 50 GB
  - Recommended use: codes, executables, data files, etc., to be used by large group of people.
  - Automounted via NFS; therefore, you need to cd into a directory to see it – it will not show up until after that.
  - If you don't have project space, for the tutorial, project space is available in /usr/projects/cesmusers and /usr/projects/workroom
  - More information: http://hpc.lanl.gov/turquoise_filesystems
- /scratch1/username and /scratch2/username
  - Recommended use: workspace for running jobs.
  - No user quotas but jobs fail when they are full (scratch1 has been near capacity lately)
  - **Scrubbed every 20 days, NO BACK-UPS. When files are deleted, they are gone, and there is nothing left but tears and recriminations.**
- Long-term storage of large data sets: use the turquoise archive system: http://hpc.lanl.gov/turquoise_archive

Transferring Files

You may use scp to transfer files to and from mapache:

If I want to transfer a file to mapache from my local machine:

scp filename username@wtrw.lanl.gov:username@mp-fe1.lanl.gov:/users/username/

Or from mapache to my local machine:

scp username@wtrw.lanl.gov:username@mp-fe1.lanl.gov:/users/username/file .

Note that you will need to authenticate with a password from your cryptocard.

Module System

- You can access multiple versions of compilers, debuggers, MPI libraries, and many other software tools via *Modulefiles*.

- Modulefiles allow you change between various version of software, so if a new bug fix in a compiler has bad consequences for your code, you can use the old compiler and keep working.

- The first step is to get a view of all the modulefiles available. Issue the command:

module avail

```
mp-fe2.lanl.gov-/users/rauscher % module avail

------------------- /usr/share/modules/modulefiles/compiler --------------------
Compilers:          intel/11.1.072     pgi/11.0
gcc/4.3.3           pgi/10.9           pgi/9.0-3(default)

------------------- /usr/share/modules/modulefiles/debugger --------------------
Debuggers:                  totalview/8.7.0-3(default)

---------------------- /usr/share/modules/modulefiles/misc ----------------------
Misc:              modules           use.user-packages
module-info        use.own

---------------------- /usr/share/modules/modulefiles/mpi -----------------------
MPI_Libraries:                openmpi-intel/1.4.3(default)
openmpi-gcc/1.4.3(default)    openmpi-pgi/1.4.3(default)

------------------- /usr/share/modules/modulefiles/packages --------------------
Package-Modules: friendly-testing hpc-tools

---------------------- /usr/share/modules/modulefiles/tools ----------------------
Tools:                      acml-intel/4.3.0-mp-int64 idl/7.0
acml-gcc/4.3.0              acml-pgi/4.3.0            lapack/3.1.1
acml-gcc/4.3.0-int64        acml-pgi/4.3.0-int64      papi/3.6.1
acml-gcc/4.3.0-mp           acml-pgi/4.3.0-mp         python/2.5
acml-gcc/4.3.0-mp-int64     acml-pgi/4.3.0-mp-int64   python/2.5.2(default)
acml-intel/4.3.0            atlas/3.8.2               subversion/1.4.2
acml-intel/4.3.0-int64      fftw/3.1.2
acml-intel/4.3.0-mp         grace/5.1.20
mp-fe2.lanl.gov-/users/rauscher %
```

Loading modules

If you want to load the default pgi compiler and mpi:

module load pgi/9.0.3
module load open-mpi/1.4.3
module load svn

To see which modules you have loaded, do

module list

If you want to unload all modules (e.g., to switch from pgi to intel compilers and start clean), do

module purge

Or just type the following to get a list of options

module

About Mapache

- Mapache is a SGI XE1300 cluster with dual-socket Intel quad-core Xeon X5550 processors. The CPUs run at 2.66 GHz. The Mapache cluster has 592 compute nodes; its aggregate performance is 50.4 TF/s with 14.2 TB of memory on 4736 cores.
- Each node has 2 Quad-Core processors for a total of 8 Intel Xeon Nehalem processor cores per node. Each node has 24 GB of memory, or 3 GB per core.

Submitting Jobs

- On the mapache front-end you log in and then submit jobs from there, both interactive and batch, which run on the back-end compute servers.

- The cluster front-end nodes are primarily for compiling building scripts, submitting jobs, editing, and simple graphics. If you try to run anything substantial on the front ends (mp-fe1 or mp-fe2) you may cause slow system response and risk memory problems for yourself and others on the system. A call from the HPC consultants would be likely.

- Mapache uses the MOAB Workload Manager
  - Jobs are submitted via the *msub* command, e.g.
    msub myjob.cmd
    where myjob.cmd is a plain text file containing your job commands
  - To check job status
    showq -w user=username
  - To kill a job
    mjobctl –c jobnumber

For more information see: http://int.lanl.gov/projects/asci/training/Moab/

Creating a Job (done automatically by CESM, but you may have to edit your run script)

```
#!/bin/tcsh
#===============================================================================
# This is a CESM Moab batch job script for mapache
#===============================================================================

#MSUB -N mycase.1850  (NAME OF JOB)
#MSUB -l nodes=2:ppn=8 (REQUEST 2 NODES with 8 PROCS EACH, 16 procs total)
#MSUB -l walltime=00:59:00 (WALL TIME)
#MSUB -A S11_CESM (IF YOU BELONG TO A SPECIFIC GROUP/QUEUE)
#MSUB -o /users/rauscher/mycase.1850/log.o (standard out files)
#MSUB -e /users/rauscher/mycase.1850/log.e (standard error files)

… (content deleted here) …

echo "`date` -- CSM EXECUTION BEGINS HERE"

module load pgi/9.0-3
module load openmpi-pgi/1.4.3

setenv OMP_NUM_THREADS 1
mpirun -np 16 ./ccsm.exe >&! ccsm.log.$LID

wait
echo "`date` -- CSM EXECUTION HAS FINISHED

cd $CASEROOT
./Tools/ccsm_postrun.csh || exit 1
```

CESM Group/Directory: Input Data and Machine-Specific Scripts

- All of you have been added to the group
- /usr/projects/cesm: contains input data, scripts, and software for CESM users (2 TB capacity)
  - scripts: machine-specific scripts (sets compilers, configures batch scripts for job submission)
  - input_data: When you build CESM, it checks for necessary input data files and downloads necessary files via SVN. File sizes can be large/slow to download; this is remedied with a common data directory.
  - software: netcdf library, visualization/processing tools (NCO, ncview, NCL, grads, CDO) (no module files available)
  - diagnostics: AMWG and other working group diagnostics scripts and input data (requires NCL and minimum 1 year model output)
  - DO NOT MAKE A WORKING DIRECTORY HERE. IT WILL BE DELETED.

Script to set up CESM environment

If you want an easy way to set up a working environment for CESM, you can source the following script /usr/projects/cesm/scripts/mapache/pgi/cesm_env_script (there is one for the intel compiler, too: /usr/projects/cesm/scripts/mapache/intel), or just add the content to your .cshrc, .tcshrc, or .login file.

```
#This script sets up the modules and environment variables for running CESM1.0.2
#with the pgi compiler
#This includes using  compatably compiled netCDF.
#This is still a work in progress.

module load pgi/9.0-3
module load openmpi-pgi/1.4.3

setenv NETCDF /usr/projects/cesm/software/conejo/netcdf
setenv INC_NETCDF /usr/projects/cesm/software/conejo/netcdf/include
setenv LIB_NETCDF /usr/projects/cesm/software/conejo/netcdf/lib
setenv INC_MPI /opt/OpenMPI/openmpi-1.4.3-pgi/include
setenv LIB_MPI /opt/OpenMPI/openmpi-1.4.3-pgi/lib64
setenv NCARG_ROOT /usr/projects/cesm/software/ncl
set path = ($path /usr/projects/cesm/software/conejo/netcdf/bin /usr/projects/cesm/software/ncl/
    bin /usr/projects/cesm/software/ncl/lib /usr/projects/cesm/software/conejo/nco/bin /usr/
    projects/cesm/software/conejo/ncview-1.93g /usr/projects/cesm/software/conejo/ImageMagick/
    bin .)
```

Getting Help

- Email consult@lanl.gov for HPC questions
- Online Documentation: http://hpc.lanl.gov